

A Methodology for Troubleshooting Inter-domain IP Multicast



Engineering Workshops

Problems Addressed

- The main types of problems addressed in this section are topology/reachability problems – the packets aren't flowing.
- The source and receiver are assumed to be in two different AS's. Troubleshooting multicast within your own campus network is a subset of interdomain troubleshooting.
- Because it is the most common today, we assume ASM. Many problems would go away with SSM.
- We will mention some things about performance issues at the end, and list some tools/references.



Why the need for a “methodology”?

- Most engineers don't troubleshoot multicast problems as often as unicast.
- As we have learned, multicast is receiver-driven (somewhat backwards).
- The problem can be far from the symptom.
- The same symptom can have many different causes, at different places in the path.



Engineering Workshops

Overview

Gather information

**Verify receiver
interest**

**Verify knowledge of
active source**

**Trace forwarding
state back**



Engineering Workshops

STEP 1: GATHER INFORMATION



Engineering Workshops

What is the problem?

Nobody can see me!

Multicast is broken ... again

Some sites can hear us, but others can't.

We're not getting anything.

Site X called to say they can't see my presentation!

Multicast isn't working between here and there.

Gather Information

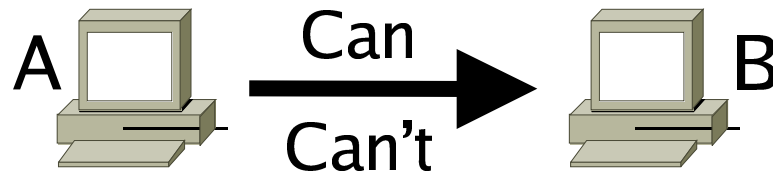
- End-users seem to have trouble reporting multicast problems in our language.
- Performance issue vs. topology/reachability issue?
- Was it working recently then stopped working, or has one site gotten nothing at all from another site?
- Is the problem intermittent, cyclic, or steady-state?
- User education about how to report a problem before a problem happens is very helpful!



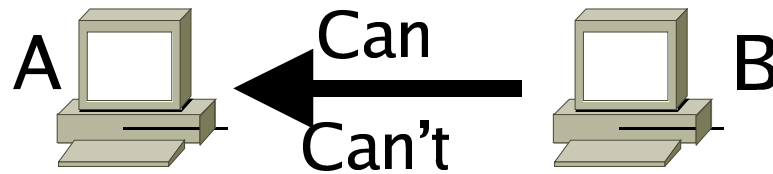
Engineering Workshops

Gather Information

- Pick ONE direction (that *is* the problem, or seems representative of the problem).
- Identify source end and receiving end.
- Recall multicast is *unidirectional* in nature...



Implies almost ***nothing*** about...



INTERNET®
2

Engineering Workshops

Gather Information

Now that you have a direction, you will need:

- A *constantly active* source IP address
- A *constantly active* receiver IP address
- The group address

It is virtually impossible to debug a multicast problem without specifying all of these!!!



Engineering Workshops

Gather Information

- OK – we know the IP addresses for the problem source, receiver, and group, and that the source and receiver are active.

Move on to step 2...



Engineering Workshops

STEP 2: VERIFY RECEIVER INTEREST



Engineering Workshops

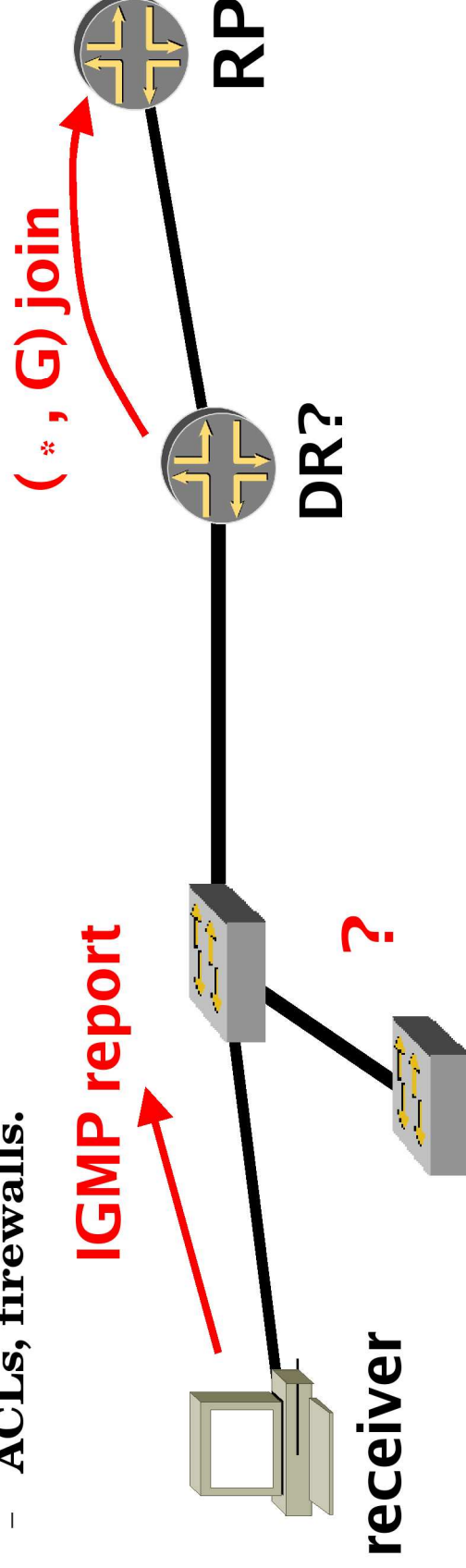
Verify Receiver Interest

- **Because of the way multicast distribution trees are built, it is almost always easier to debug a problem by starting at the receiver. If you are the sender, you are pretty much working blind.**
- **Recall in ASM, group interest on a subnet is indicated by a host sending out (multicast) an IGMPv2 membership report.**
- **The DR (designated router) on a segment is responsible for listening to that report, and forwarding a PIM (*, G) join towards the RP.**
- *For this step, all we need to do is verify which router is the DR, and check that it knows it has interested listeners for that group on the interface facing the receiver. Stop there. Don't worry about getting to the RP at this point.*

Verify Receiver Interest

What can go wrong?

- No host is sending out IGMP membership reports, or not the right version.
- A switch is in the path that is dropping/limiting multicast/IGMP.
- The router is not running IGMP, PIM, etc.
- A device has been elected DR that shouldn't have been.
- bugs, incompatible timer implementations, querier confusion, etc.
- ACLs, firewalls.



DR? Gack! I dunno where RP...

Verify Receiver Interest

- You might think you know which router is the DR, but you should not proceed until it has been verified. It only takes a couple seconds.
- To verify the DR, log into the router you think *should* be routing multicast for the receiver.
 - 1) Find/verify the interface that serves the receiver's subnet.
 - 2) Check that there is no other PIM router that thinks *it* is the DR for the subnet.
- Although in our workshop lab our first-hop routers are Ciscos, the following examples show both Junipers and Ciscos.



Verify Receiver Interest

1) Cisco: find the right interface:

 receiver

```
squash# show ip rpf 140.221.34.1
```

```
RPF information for ws-video.mcs.anl.gov (140.221.34.1)
```

```
RPF interface: GigabitEthernet5/7
```

```
RPF neighbor: ? (0.0.0.0) - directly connected
```

```
RPF route/mask: 140.221.34.0/28
```

```
RPF type: unicast (connected)
```

```
RPF recursion count: 0
```

```
Doing distance-preferred lookups across tables
```

```
squash#
```



Engineering Workshops

Verify Receiver Interest

1) Juniper: find the right interface:

receiver

```
remote@MREN-M5> show multicast rpf 140.221.34.1  
Multicast RPF table: inet.2, 5051 entries
```

```
140.221.34.0/27
```

```
Protocol: Direct
```

```
Interface: ge-0/0/0.108
```



Engineering Workshops

Verify Receiver Interest

1) Nortel: find the right interface:

```
ROUT01:3# show ip igmp group
```

Igmp Group

GRPADDR	INPORT	MEMBER	EXPIRATION	TYPE
224.0.1.22	2/5	142.90.105.25	1257	Dynamic



receiver



Engineering Workshops

Verify Receiver Interest

2) Cisco: verify DR for that interface:

```
squash#sh ip igmp interface gig5/7
GigabitEthernet5/7 is up, line protocol is up
Internet address is 140.221.34.13/28
IGMP is enabled on interface
Current IGMP host version is 2
Current IGMP router version is 2
IGMP query interval is 60 seconds
IGMP querier timeout is 120 seconds
IGMP max query response time is 10 seconds
Last member query response interval is 1000 ms
Inbound IGMP access group is not set
IGMP activity: 867 joins, 866 leaves
Multicast routing is enabled on interface
Multicast TTL threshold is 0
Multicast designated router (DR) is 140.221.34.13 (this system)
IGMP querying router is 140.221.34.13 (this system)
No multicast groups joined
squash#
```



Verify Receiver Interest

2) Juniper: verify DR for that interface:

```
remote@MREN-M5> show pim interfaces
```

Instance: PIM.master

Name	Stat	Mode	IP	V	<u>State</u>	Count	<u>DR address</u>
at-0/2/1.237	Up	Sparse	4	2	P2P	1	
at-0/2/1.6325	Up	Sparse	4	2	P2P	1	
at-0/2/1.9149	Up	Sparse	4	2	P2P	1	
ge-0/0/0.108	Up	Sparse	4	2	DR	1	140.221.34.13
ge-0/0/0.109	Up	Sparse	4	2	NotDR	1	10.10.10.1

```
remote@MREN-M5>
```



Engineering Workshops

Verify Receiver Interest

2) Nortel: verify DR for that interface:

```
ROUT01:3# show ip pim interface
```

Pim Interface

IF	ADDR	MASK	MODE	DR
Port2/1	142.231.1.50	255.255.255.248	sparse	142.231.1.54
Port2/3	206.12.24.150	255.255.252.0	sparse	0.0.0.0

```
ROUT01:3#
```



Engineering Workshops

Verify Receiver Interest

- **SO... now you are sure you are on your receiver's DR.**
- **Remember, multicast is receiver-driven.**
- **QUESTION: Does the DR know that there are interested receivers of the group on your host's subnet??**
- **Look at IGMP for the group in question.**



Verify Receiver Interest

On the DR (Cisco):

**group you are
debugging**



```
squash#sh ip igmp group 233.2.171.1
IGMP Connected Group Membership
Group Address Interface      Uptime    Expires    Last Reporter
233.2.171.1   Vlan1      1d03h     00:02:16   140.221.10.87
233.2.171.1   GigabitEthernet5/7 7w0d      00:02:21   140.221.34.1
squash#
```

If receiver's interface is in this list, you are OK. You might want to watch for a while to ensure no timeouts are occurring.



Engineering Workshops

Verify Receiver Interest

On the DR (Juniper):

**group you are
debugging**



```
remote@MREN-M5> show igmp group 233.2.171.1
Interface: ge-0/0/0.108
  Group: 233.2.171.1
    Source: 0.0.0.0    Last Reported by: 206.220.240.86
    Timeout:    156      Type: Dynamic
remote@MREN-M5>
```

**If receiver's interface is in this list, you are OK.
You might want to watch for a while to ensure
no timeouts are occurring.**



Engineering Workshops

Verify Receiver Interest

On the DR (Nortel):

```
ROUT01:3# show ip igmp group
```

Igmp Group

GRPADDR	INPORT	MEMBER	EXPIRATION	TYPE
224.0.1.22	2/5	142.90.105.25	1202	Dynamic
224.0.1.22	2/5	142.90.106.30	1197	Dynamic

If receiver's interface is in the list for the group, you are OK.

You might want to watch for a while to ensure no timeouts are occurring.



Engineering Workshops

Verify Receiver Interest

- **What if your interface isn't listed with that group, even though everything else about the DR looked fine??**
- **You have a problem!**
 - **Host OS / driver problem**
 - **Application problem**
 - **Broken IGMP snooping switches in the middle**
 - **Try tcpdump on the host - can you see the IGMP membership reports on the wire? (Remember, they don't have to come from that particular host.)**



Verify Receiver Interest

- **If your receiver's DR knows it has listeners of your group on that interface, you are done this step.**

Move on to step 3...



Engineering Workshops

STEP 3: VERIFY KNOWLEDGE OF ACTIVE SOURCE



Engineering Workshops

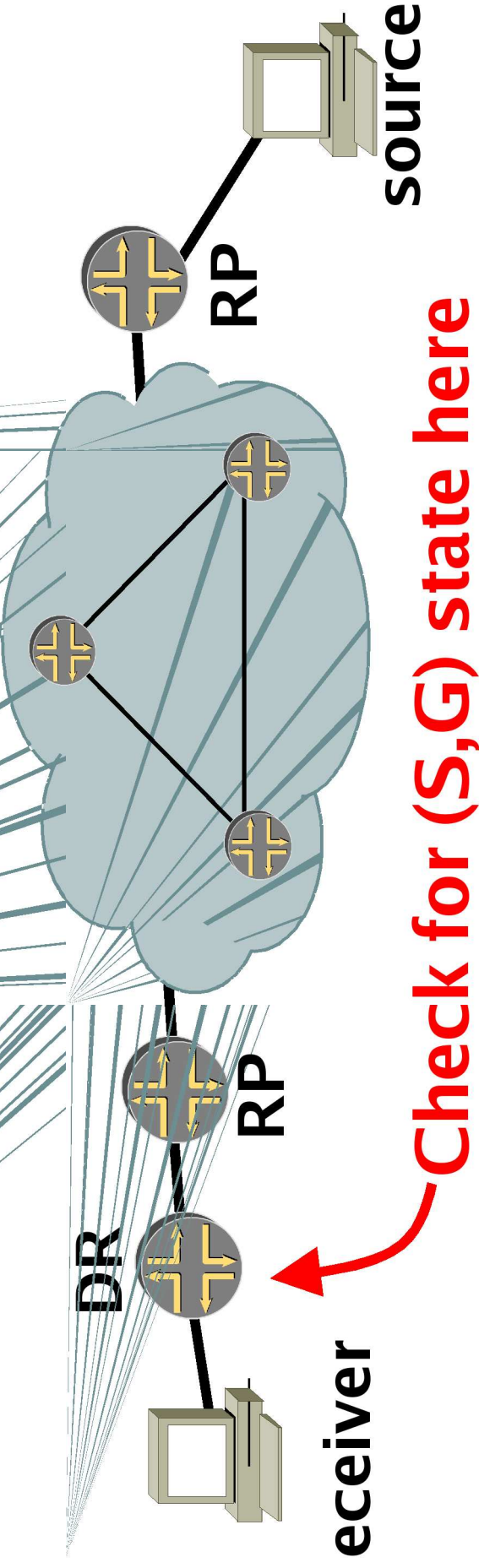
Verify knowledge of active source

- **This is often the most complex part – the bulk of your work could be here. As we have learned, a lot has to happen for the receiver's DR to know about a particular source.**
- **You MAY have view this from both ends**
 - **The receiver's RP**
 - **The source's RP**
- **For most interdomain cases, these RPs will not be the same, and MSDP will be involved.**

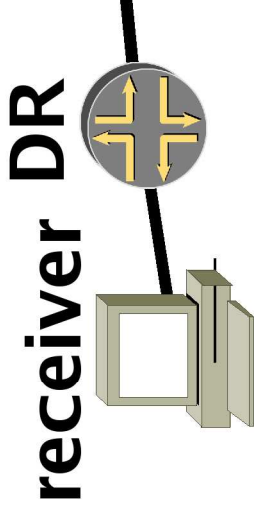


Verify knowledge of active source

- First, let's check to see if this is a problem at all.
- If the receiver's DR has (S,G) state already, we know we are ok on knowledge of active source, and we can skip this whole step!



Verify knowledge of active source



On the receiver's DR (Cisco):

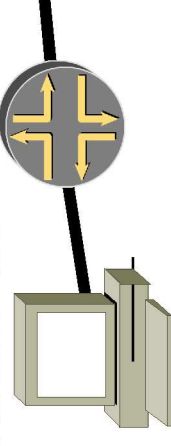
```
squash# show ip mroute 233.2.171.1 141.142.64.104
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, s - SSM Group, C - Connected, L - Local,
P - Pruned, R - RP-bit set, F - Register flag, T - SPT-bit set,
J - Join SPT, M - MSDP creat entry, X - Proxy Join Timer Running
A - Advertised via MSDP, U - URD,
I - Received Source Specific Host Report
Outgoing interface flags: H - Hardware switched
Timers: Uptime/Expires
Interface state: Interface, Next-Hop or VCD, State/Mode
(141.142.64.104, 233.2.171.1), 1w0d/00:02:59, flags: CJT
Incoming interface: Vlan669, RPF nbr 130.202.222.74 GOOD!
Outgoing interface list:
GigabitEthernet5/7, Forward/Sparse, 20:19:14/00:02:08
Vlan1, Forward/Sparse, 1w0d/00:01:56
```



Engineering Workshops

Verity knowledge of active source

receiver DR



On the receiver's DR (Juniper):

```
remote@starlight-m10> show multicast route group 233.2.171.1  
source-prefix 141.142.64.104
```

Family:	INET	Source prefix	Act	Pru	InIf	NHid	Session Name
Group		141.142.64.104 /32	A	F	6	246	Static Alloc

GOOD!

...extensive)

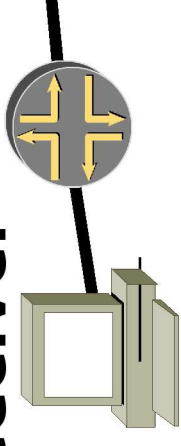
Family:	INET	Source prefix	Act	Pru	NHid	Packets	IfMi	Timeout
Group		233.2.171.1 141.142.64.104 /32	A	F	246	8702556	69	360
Upstream interface: ge-0/0/0.0								
Session name: Static Allocations								
Forwarding rate: 1 kBps (9 pps)								



Engineering Workshops

Verify knowledge of active source

receiver DR



On the receiver's DR (Nortel):

```
ROUT01:3# show ip pim mroute src 131.188.3.221 grp
224.0.1.1
Src: 131.188.3.221 Grp: 224.0.1.1 RP: 207.23.240.200Upstream:
142.231.1.54
Flags: SPT CACHE SG
Incoming Port: Port2/1 ,
Outgoing Ports: Vlan1-2/5,
```

GOOD!



Engineering Workshops

Verify knowledge of active source

- If the DR does NOT know about the source, we may only see a (*, G) entry on a Cisco DR, and we have some work to do.

```
squash# show ip mroute 233.2.171.1 141.142.64.104
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, s - SSM Group, C - Connected, L - Local,
       P - Pruned, R - RP-bit set, F - Register flag, T - SPT-bit set,
       J - Join SPT, M - MSDP creat entry, X - Proxy Join Timer Running
       A - Advertised via MSDP, U - URD,
       I - Received Source Specific Host Report
Outgoing interface flags: H - Hardware switched
Timers: Uptime/Expires
Interface state: Interface, Next-Hop or VCD, State/Mode
(*, 233.2.171.1), 7w0d/00:02:59, RP 192.5.170.2, flags: SJCF
Incoming interface: Vlan29, RPF nbr 140.221.20.97
Outgoing interface list:
  GigabitEthernet5/7, Forward/Sparse, 20:22:27/00:02:52
  Vlan1, Forward/Sparse, 7w0d/00:02:45
```

(* , G) only is BAD!



Engineering Workshops

Verify knowledge of active source

- If the DR does **NOT** know about the source, we may see nothing on a Juniper DR, and we have some work to do.

```
remote@starlight-m10> show multicast route group 233.2.171.1
                               source-prefix 141.142.64.104
Family: INET
Group          Source prefix      Act Pru InIf  NHid  Session Name
remote@starlight-m10>
```

BAD!



Engineering Workshops

Verify knowledge of active source

- If the DR does **NOT** know about the source, we may see nothing on a Nortel DR, and we have some work to do.

```
ROUT01:3# show ip pim mroute grp 233.2.171.1 src 233.1.15.16
```

```
Pim Multicast Route  
Total Num of Entries Displayed 0
```

```
ROUT01:3#
```

BAD!



Engineering Workshops

Verify knowledge of active source

1

- **Recall that knowledge of active sources is first spread through a given PIM domain by per-group RP-rooted shared distribution trees.**
- **Current practice is to set the Source Path Tree (SPT) threshold to zero, so that (S,G) state is created by on the first packet sent through the RP.**
- **But if the RPT doesn't get built properly, the SPT never will!**



Engineering Workshops

Verify knowledge of active source

1

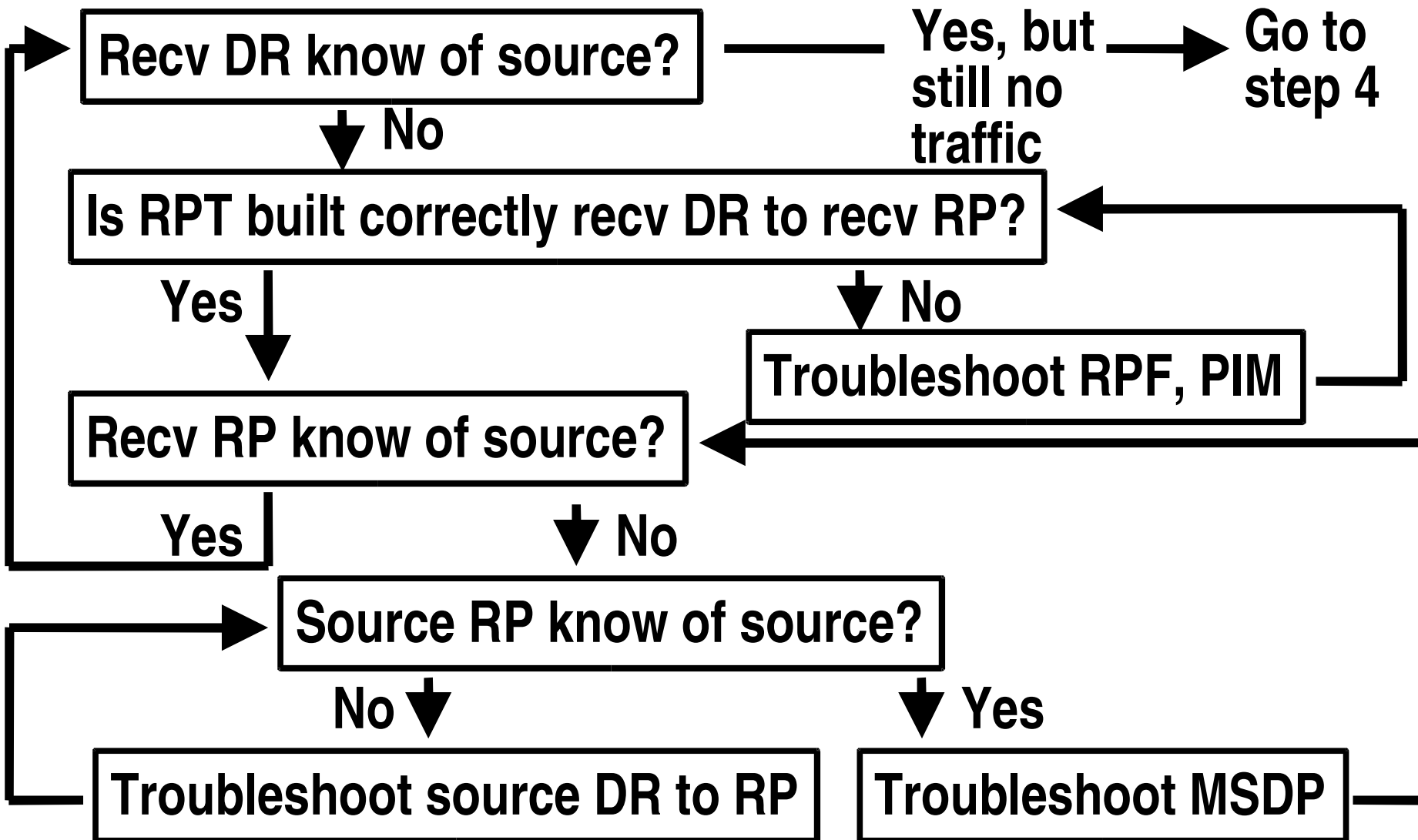
- **So, first, we will work back from the receiver's DR to its RP, to be sure that the RPT branch is built correctly.**
- **Second, we will check to see if the receiver's RP knows about the source.**
- **Third, we will check with the source end for their RP's knowledge and advertisement of the source.**
- **Last, we will troubleshoot MSDP as needed to make sure knowledge of the source can get from one RP to the other.**
- **The following page has a rough flowchart for later reference.**



Engineering Workshops

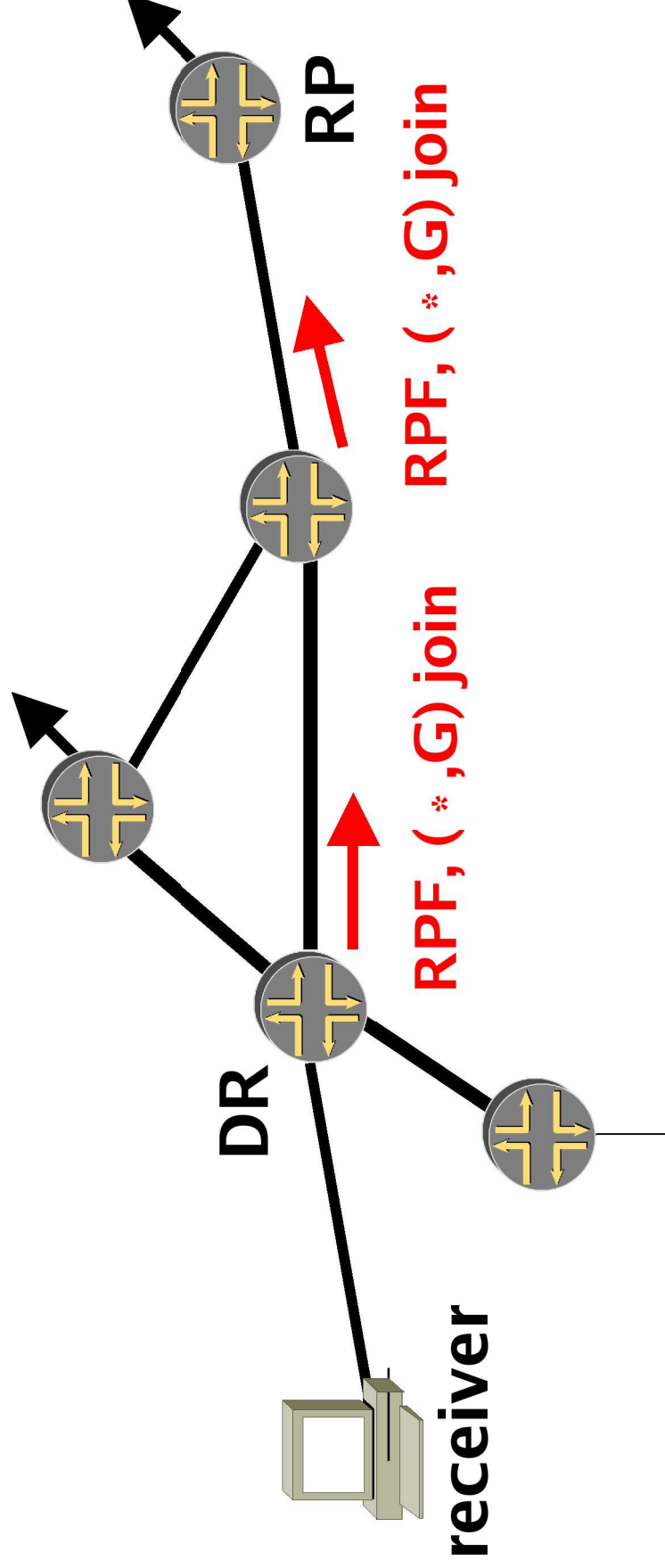
Verify knowledge of active source

1



Verify knowledge of active source

- First, we check that the RPT is built properly from the receiver's DR back to the receiver's RP.



Verify knowledge of active source

- **Does the DR have the right RP (Cisco)?**
 - We can first just look at the (*, G) entry on the receiver's DR.
 - If that doesn't look right, we can look to see how it learned about the RP with `show ip pim rp mapping <group>` .

```
squash# show ip mroute 233.2.171.1
IP Multicast Routing Table
Flags: D - Dense, S - Sparse, s - SSM Group, C - Connected, L - Local,
       P - Pruned, R - RP-bit set, F - Register flag, T - SPT-bit set,
       J - Join SPT, M - MSDP creat entry, X - Proxy Join Timer Running
       A - Advertised via MSDP, U - URD,
       I - Received Source Specific Host Report
Outgoing interface flags: H - Hardware switched
Timers: Uptime/Expires
Interface state: Interface, Next-Hop or VCD, State/Mode
(*, 233.2.171.1), 7w0d/00:02:59, RP 192.5.170.2, flags: SJCF
Incoming interface: Vlan29, RPF nbr 140.221.20.97
Outgoing interface list:
  GigabitEthernet5/7, Forward/Sparse, 20:22:27/00:02:52
  Vlan1, Forward/Sparse, 7w0d/00:02:45
```



Verify knowledge of active source

1

- Does the DR have the right RP (Juniper)?

```
remote@MREN-M5> show pim rps detail
```

```
Instance: PIM.master
```

```
Family: INET
```

```
RP: 206.220.241.254
```

```
Learned via: static configuration
```

```
Time Active: 13w2d 09:59:40
```

```
Holdtime: 0
```

```
Group Ranges:
```

```
224.0.0.0/4
```

```
Active groups using RP:
```

```
224.2.127.254
```

```
233.2.171.1
```

```
239.22.33.5
```

```
total 3 groups active
```

```
remote@MREN-M5>
```



Engineering Workshops

Verify knowledge of active source

- **Does the DR have the right RP (Nortel)?**

```
ROUT01:3# show ip pim active-rp 224.255.222.23
```

Pim Grp->RP Active RP Table

GRPADDR	RP-ADDR	RP-PRIORITY
224.255.222.239	207.23.240.200	0

```
ROUT01:3#
```



Verify knowledge of active source ¹

- **What if the RP is wrong?**
 - **A common problem is that auto-RP and/or PIMv2 BSR may be running without the admin's knowledge (on Ciscos they are on by default when PIM-SM is enabled, and Junipers listen to them). Information can leak from a neighboring AS! These take precedence over anything you statically configure. Hint: use `ip pim rp-address <address> override`**
 - **Auto-RP and BSR are complex, and could have any one of a number of problems. We recommend static configuration in most campus networks, Anycast-RP in backbone/transit networks.**
 - **Might just be a typo in entering the static RP address.**



Verify knowledge of active source

1

- **Now that you are sure of what the RP is (and it is correct), starting at the receiver's DR, work your way back to the receiver's RP:**
- **Check that the RPF is pointing the way you expect.**
- **Check that PIM is configured and working properly on the interface. A common problem is PIM is not turned on for a particular interface.**
- **You may also want to double-check that each router has (*, G) state for the group you are debugging.**



Verify knowledge of active source

1

Cisco:

- show ip rpf <RP ip address>
- show ip pim neighbor <rpf interface>

```
squash# show ip rpf 192.5.170.2
RPF information for kiwi-loop.anchor.anl.gov
(192.5.170.2)
  RPF interface: Vlan29
  RPF neighbor: kiwi.anchor.anl.gov (140.221.20.97)
  RPF route/mask: 192.5.170.2/32
  RPF type: unicast (ospf 683)
  RPF recursion count: 0
  Doing distance-preferred lookups across tables
```

```
squash# show ip pim neighbor Vlan29
PIM Neighbor Table
Neighbor Address    Interface    Uptime Expires    Ver    Mode
140.221.20.97      Vlan29      7w0d  00:01:35    v2     (DR)
squash#
```



Engineering Workshops

Verify knowledge of active source

1

Juniper:

- show multicast rpf <RP ip address>
- show pim neighbors

```
remote@MREN-M5> show multicast rpf 206.220.241.254
```

Multicast RPF table: inet.2, 5061 entries

206.220.241.0/24

Protocol: BGP

Interface: ge-0/0/0.108

```
remote@MREN-M5> show pim neighbors
```

Instance: PIM.master

Interface	IP	V	Mode	Option	Uptime	Neighbor addr
at-0/2/1.237	4	2		H	4w6d11h	192.122.182.13
at-0/2/1.6325	4	2		H	4w6d11h	206.166.9.33
at-0/2/1.9149	4	2		HP B	4w6d11h	199.104.137.245
ge-0/0/0.108	4	2		H G	4w6d11h	140.221.20.97



Engineering Workshops

Verify knowledge of active source

1

Nortel:

- show ip pim neighbor

```
ROUT01:3# show ip pim neighbor
```

Pim Neighbor

INTERFACE	ADDRESS	UPTIME	EXPIRE
Port2/1	142.231.1.54	20 day(s),	08:25:11 0 day(s)

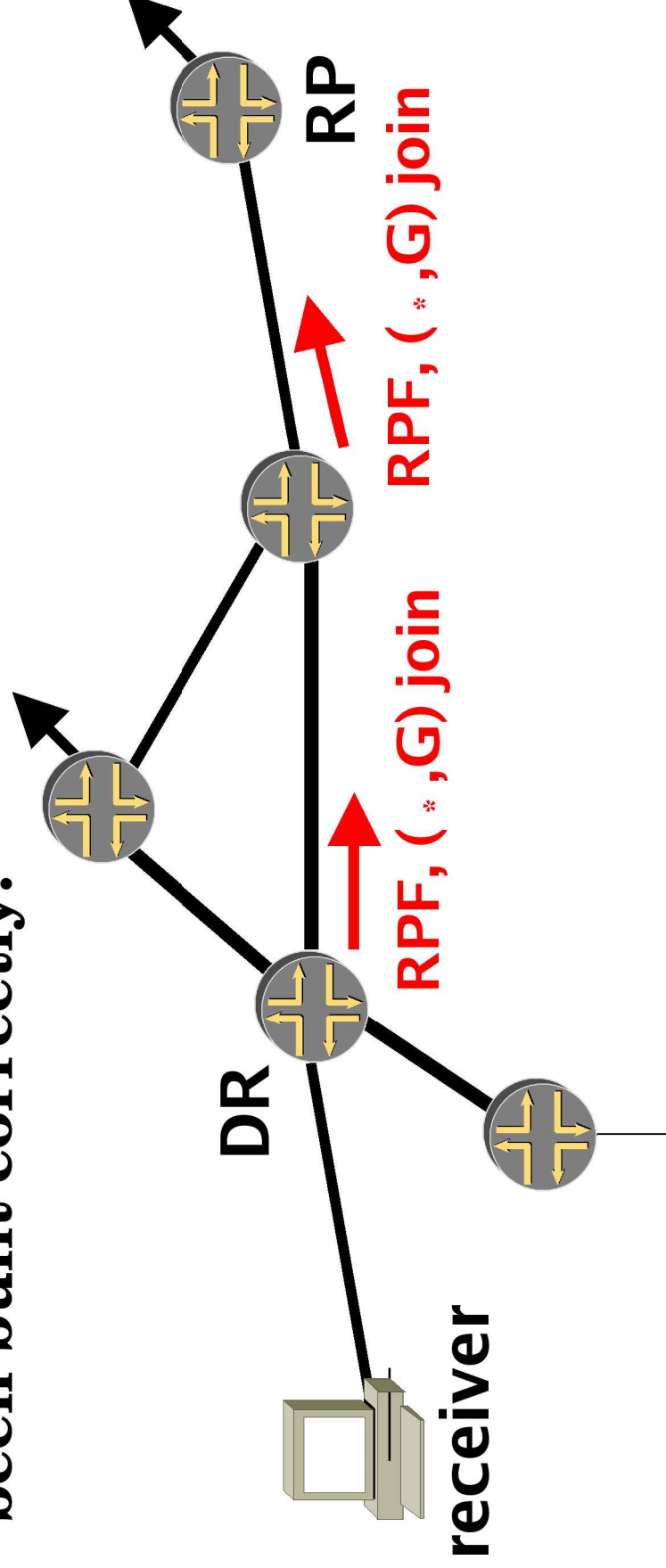


INTERNET®

Engineering Workshops

Verify knowledge of active source

- Repeat that process until you have verified the RPF paths and the PIM adjacencies back to the receiver's RP. This verifies that the RPT has been built correctly.



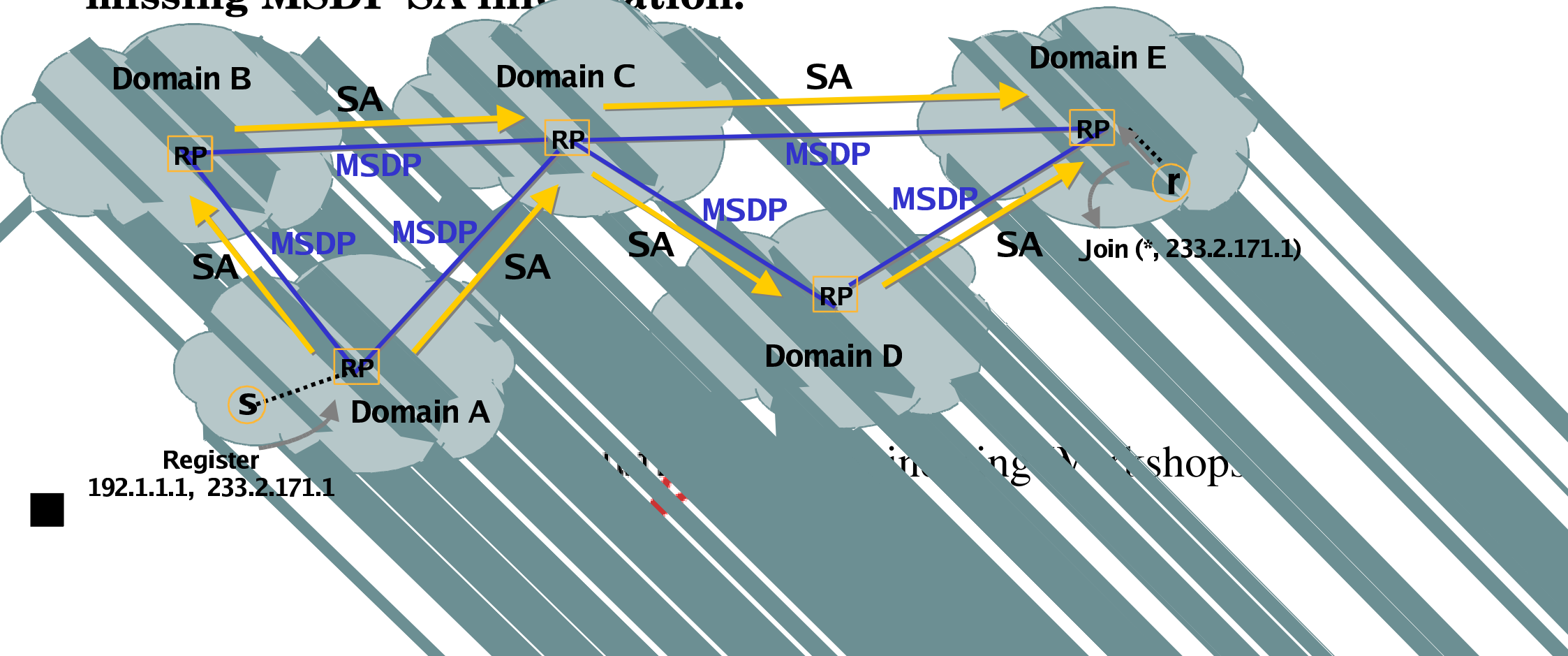
Verify knowledge of active source

- **Next Big Question: Does the receiver's RP have knowledge of the active source?**
- **Since we already checked that the RPT is correct, it probably doesn't, or the DR would have likely had (S,G) information.**
- **If it doesn't, but has (*, G) only, and no MSDP SA (source-active) cache entry for that source, we will have to find out some information about the source end of things, then troubleshoot MSDP.**
- **Note it does not matter which peer you get an SA from as long as it is accepted and in the cache. However, if you are going to open a ticket with an upstream, you might as well figure out who you expect to accept it from.**



Verify knowledge of active source

- The objective here will be to get an MSDP source-active about the source to our receiver's RP.
- The SA originates *from the source's RP*, and is re-advertised/ flooded by MSDP peers along the way.
- Some sites have estimated that about half of their multicast problems are problems associated with missing MSDP SA information.



Verify knowledge of active source

1

On the receiver's RP:

```
Kiwi#sh ip mroute 233.2.171.1 141.142.64.102
```

IP Multicast Routing Table

Flags: D - Dense, S - Sparse, B - Bidir Grp, s - SSM Grp, C-Connected,
L - Local, P - Pruned, R - RP-bit set, F - Register flag,
T - SPT-bit set, J - Join SPT, M - MSDP created entry,
X - Proxy Join Timer Running, A - Candidate for MSDP Advert,
U - URD, I - Recved Source Specific Host Rpt, Z - Mcast Tunnel,
Y - Joined MDT-data group, y - Sending to MDT-data group

Outgoing interface flags: H - Hardware switched

Timers: Uptime/Expires

Interface state: Interface, Next-Hop or VCD, State/Mode

(* , 233.2.171.1), 6w6d/stopped, RP 192.5.170.2, flags: S

Incoming interface: Null, RPF nbr 0.0.0.0

Outgoing interface list:

GigabitEthernet5/0, Forward/Sparse, 6w6d/00:03:01

BAD!

```
Kiwi#sh ip msdp sa-cache 233.2.171.1 141.142.64.102
```

MSDP Source-Active Cache

Entry not found

BAD!



Engineering Workshops

Verify knowledge of active source

1

- **Recall it is MSDP's job to flood source-active advertisements between peers so that an RP in one PIM domain can know about active sources in another.**
- **MSDP SA advertisements are accepted/forwarded or rejected based on MSDP "peer-RPF" rules covered earlier in this workshop.**
- **Remember, the information being tested against the peer-RPF rules is the originating RP's IP address. Not the IP of the source itself, but its RP.**
- **We need to trace the source-RP via the peer-RPF rules from our receiver's RP out into our neighbor's AS.**



Engineering Workshops

Verify knowledge of active source

- **But... how do we know the source's RP if we run only the receiver network?**
 - **You may have to pick up phone and walk them through verifying the source's DR and finding the group-to-RP mapping there.**
 - **Get them to tell you they have verified the source is sending, and the IP of their RP is ____.**
 - **You might want to have them look to see that they mark the mroute as a candidate for MSDP advertisement while you're there. (Example - next slide.)**

Verify knowledge of active source

1

On the *source's* RP to show generating an SA:

Source IP

```
Kiwi#sh ip mroute 233.2.171.1 141.142.64.104
IP Multicast Routing Table
Flags: D-Dense, S-Sparse, B-BidirGroup, s-SSM Group, C-Connected,
       L - Local, P - Pruned, R - RP-bit set, F-Register flag,
       T - SPT-bit set, J - Join SPT, M - MSDP created entry,
       X - Proxy Join Timer Running,
       A - Candidate for MSDP Advertisement, U - URD,
       I - Recv Source Specific Host Report, Z - Multicast Tunnel,
       Y - Joined MDT-data group, y - Sending to MDT-data group
Outgoing interface flags: H - Hardware switched
Timers: Uptime/Expires
Interface state: Interface, Next-Hop or VCD, State/Mode

(141.142.64.104, 233.2.171.1), 6w6d/00:03:26, flags: TA
  Incoming interface: GigabitEthernet5/0, RPF nbr 141.142.20.124
  Outgoing interface list:
    ATM3/0.6200, Forward/Sparse, 2w0d/00:02:42 (ttl-threshold 32)
Kiwi#
```



Engineering Workshops

Verify knowledge of active source

- **Now we have the source/originating RP's IP address.**
- **The idea here is we are trying to figure out which of our MSDP peers we should expect to get knowledge of the actual source from.**
 - **If the source RP is an MSDP peer of our RP, the source RP is the RPF peer.**
 - **If we look at `show ip mbgp <source RP IP>`, the MSDP peer in the adjacent AS is the RPF peer.**
 - **In practice, in most campus networks, `show ip rpf <source RP IP>` and `show ip mbgp <source RP IP>` will usually get you going in the right direction.**



Verify knowledge of active source

1

```
guava#sh ip rpf 141.142.20.124
```

```
RPF information for lsd6509.sl.startap.net (206.220.241.254)
```

```
RPF interface: Vlan109
```

```
RPF neighbor: mren-anl-gige.anchor.anl.gov (192.5.170.214)
```

```
RPF route/mask: 141.142.0.0/16
```

```
RPF type: mbgp
```

```
RPF recursion count: 0
```

```
Doing distance-preferred lookups across tables
```

Source's RP



```
guava#sh ip mbgp 141.142.20.124
```

```
BGP routing table entry for 141.142.0.0/16, version 1977637
```

```
Paths: (2 available, best #1, table NULL) Flag: 0x208
```

```
Advertised to peer-groups:
```

```
imbgp-mesh
```

```
22335 11537 1224
```

```
192.5.170.214 from 192.5.170.214 (206.220.241.254)
```

```
Origin IGP, localpref 40100, valid, external, best
```

```
Community: 683:65001 11537:950 22335:11537
```

```
293 11537 1224
```

```
192.5.170.78 from 192.5.170.78 (134.55.29.97)
```

```
Origin IGP, metric 100, localpref 10000, valid, external
```

```
Community: 293:52 683:293 no-export
```



Engineering Workshops

Verify knowledge of active source

1

- Assuming we do not have an entry for the source and group in our receiver RP's SA-cache, we *might* be able to see if we are getting a reasonable SA advertisement but rejecting it:

```
LSD6509#sh ip msdp sa-cache 233.2.171.1 141.142.64.104 rejected
detail read-only
MSDP Rejected SA Cache
5285 rejected SAs received over 00:00:13, cache size: 2000 entries
Timestamp (source, group)
3928782.016, (141.142.64.104, 233.2.171.1), RP: 141.142.12.1, Peer:206.220.240.220
Reason: rpf-fail
3928782.076, (141.142.64.104, 233.2.171.1), RP: 141.142.12.1, Peer:205.189.32.74
Reason: rpf-fail
3928782.120, (141.142.64.104, 233.2.171.1), RP: 141.142.12.1, Peer:205.189.32.70
Reason: rpf-fail
3928782.148, (141.142.64.104, 233.2.171.1), RP: 141.142.12.1, Peer:205.213.117.13
Reason: rpf-fail
```

This is a circular buffer, so it's hit-or-miss...

- On a Juniper, turn on MSDP traceoptions and search the file.
flag source-active receive detail



Engineering Workshops

Verify knowledge of active source

1

- If we *are* getting an SA from what we think should be the RPF peer, yet rejecting it, we need to work through the MSDP peer-RPF rules to figure out why.

Possible reasons:

- We've configured to use only the multicast RIB, yet we have no MBGP route to the originating RP. Check that the source network is advertising the route to the RP in MBGP and we are accepting it (policy misconfigurations).
- We have MBGP running, but not MSDP, with a peer that appears to have a better route to the originating RP than who we think is the RPF peer.
- incorrectly configured default peer.
- bugs, voodoo, who knows!



Engineering Workshops

Verify knowledge of active source

- **Assuming you are not getting an SA from the peer you think should be the RPF peer, you may need to open a ticket with your upstream provider or peer. You can give them the following:**
 - **We are not getting an SA for <source IP address>**
 - **The group address is <group address>**
 - **The source's RP is <source RP IP address>**
 - **We expected to get this from <MSDP peer's IP address>**
- **Also report if you are not getting the MBGP route.**

Verify knowledge of active source

1

- **Other than just turning the problem over to your upstream provider, for many Internet2 campuses, Abilene core routers will be in the path.**
- **It is sometimes helpful to go to the router proxy closest to the source and check for the SA-cache entry for the source/group in question there.**
- **If there is no entry there, it is not too surprising your campus is not getting a valid SA. (We have a screenshot at the end of these slides.)**
<http://loadrunner.uits.iu.edu/%7Erouterproxy/abilene/>



Engineering Workshops

Verify knowledge of active source ¹

- **Since you have already checked your path back from the receiver to your RP, you should then get (S,G) state on the receiver's DR when you fix rejecting a received SA, or your upstream provider or peer resolves the ticket concerning a missing SA.**

Move on to step 4...



Engineering Workshops

Overview Refresher!

Gather information

**Verify receiver
interest**

**Verify knowledge of
active source**

**Trace forwarding
state back**



Engineering Workshops

STEP 4: TRACE FORWARDING STATE BACK




Trace forwarding state back

1

- We now have (S,G) state on the receiver's DR.
- Next, we need to check to see if traffic is actually flowing...
(Cisco example)

```
squash# show ip mroute 233.2.171.1 141.142.64.104 count
IP Multicast Statistics
226 routes using 103842 bytes of memory
42 groups, 4.38 average sources per group
Forwarding Counts: Pkt Count/Pkts per second/Avg PktSize/Kilobits per sec
Other counts: Total/RPF fail/Other drops(OIF-null,rate-limit,etc)

Group: 233.2.171.1, Source count: 100, Group pkt count: 987910557
  Source: 141.142.64.104/32, Forwarding: 0/0/0/0, Other: 6/0/6
squash#
```



If this is zero, you still have a problem.



Engineering Workshops

Trace forwarding state back

1

- Here's how to check if traffic is flowing on a Juniper:

```
litvanyi@starlight-m10> show multicast route group 233.2.171.1
                           source-prefix 141.142.64.104 extensive
```

Family: INET

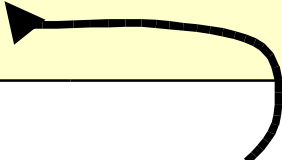
Group	Source prefix	Act	Pru	NHid	Packets	...	
233.2.171.1	141.142.64.104	/32	A	F	426	0	249

Upstream interface: ge-0/0/0.11537

Session name: Static Allocations

Forwarding rate: 0 kBps (0 pps)

```
litvanyi@starlight-m10>
```



If this is zero, you still have a problem.



Engineering Workshops

Trace forwarding state back

1

- Start on your receiver's DR.
- This time, RPF back towards the actual source IP address (as opposed to the source RP).

On a Cisco:

 **source**

```
squash# show ip rpf 141.142.64.104
RPF information for ag-nl-video.ncsa.uiuc.edu (141.142.64.104)
  RPF interface: Vlan669
  RPF neighbor: guava-stardust.anchor.anl.gov (130.202.222.74)
  RPF route/mask: 0.0.0.0/0
  RPF type: unicast (ospf 683)
  RPF recursion count: 0
  Doing distance-preferred lookups across tables
```



Engineering Workshops

Trace forwarding state back

1

On a Juniper:

source
←

```
litvanyi@starlight-m10> show multicast rpf 141.142.64.104
Multicast RPF table: inet.2, 5060 entries

204.121.50.0/24
  Protocol: BGP
  Interface: ge-0/0/0.293
  Neighbor: 198.125.140.97

litvanyi@starlight-m10>
```

- You are looking to see how you are expecting the SPT tree to be built, where you actually expect the packet flow to come from.



Engineering Workshops

Trace forwarding state back

- Work your way back towards the *source IP*, looking for PIM problems along the way.

Cisco:

```
squash# show ip pim neighbor Vlan669
```

PIM Neighbor Table

Neighbor Address	Interface	Uptime	Expires	Ver	Mode
130.202.222.74	Vlan669	7w0d	00:01:35	v2	(DR)

Juniper:

```
litvanyi@starlight-m10> show pim neighbors detail | find "ge-0/0/0.293"
```

Interface: **ge-0/0/0.293**

Address: **198.125.140.97, IPv4, PIM v2**

Hello Option Holdtime: 105 seconds 98 remaining

Hello Option DR Priority: 1

Hello Option LAN Prune Delay: delay 500 ms override 2000 ms

Rx Join: Group	Source	Timeout
233.2.171.1	203.255.248.51	201
233.2.171.1	150.183.121.105	201
233.2.171.1	131.94.133.48	201

INTERNET.

Engineering Workshops

Trace forwarding state back

1

- Log into that upstream router and check state there with:

```
router# show ip mroute <group> <source>
```

```
router# show ip mroute <group> <source> count
```

- **Or (Juniper):**

```
router> show multicast route group <group> source  
                <source> extensive
```

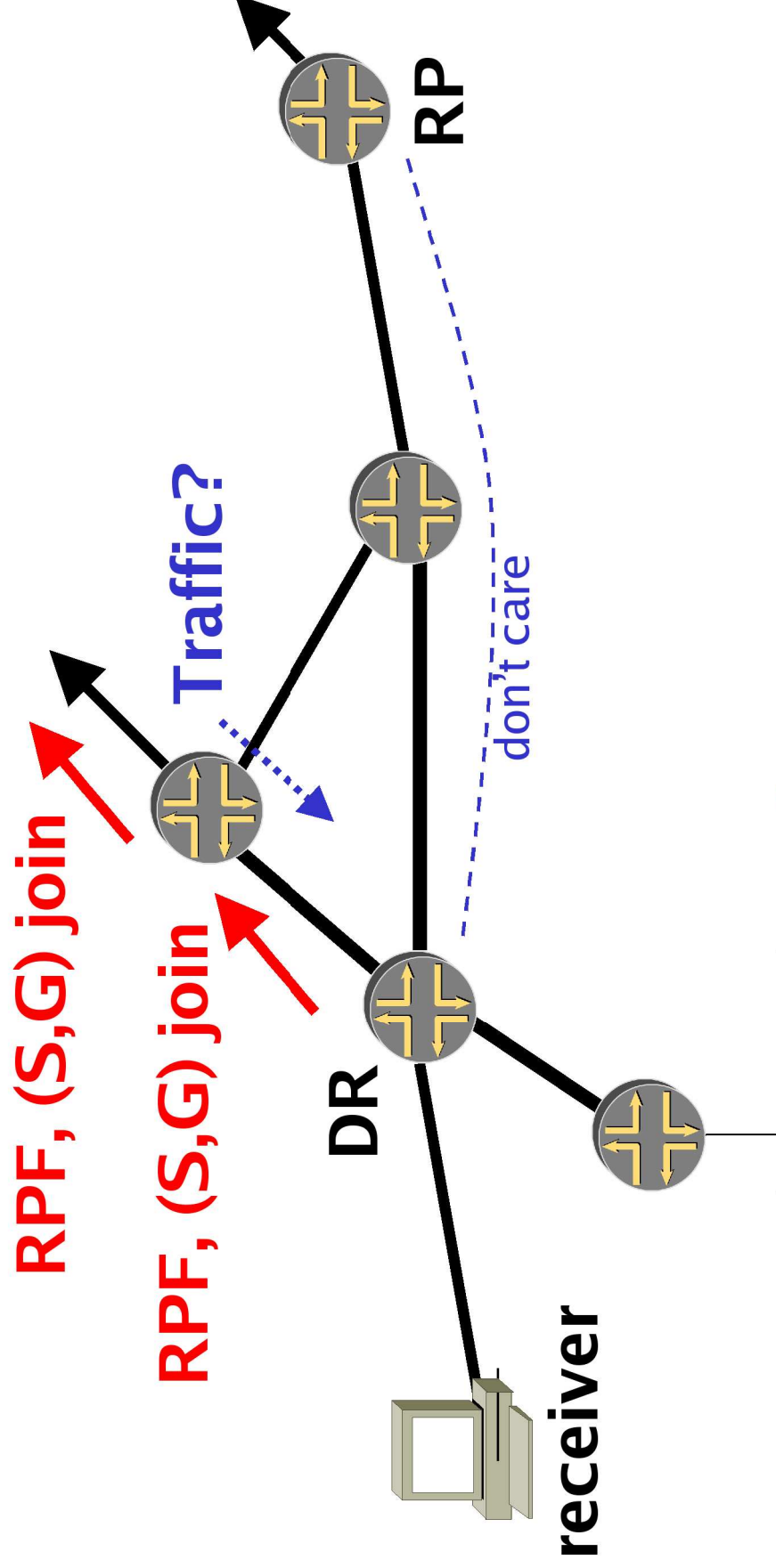
- Look to see if the downstream router is in the outgoing interface list, and to see if you see a positive traffic rate.
- Hopefully you will work your way back to a router that is seeing the traffic flow.



Engineering Workshops

Trace forwarding state back

We are tracing back the SPT...



Trace forwarding state back

1

```
Kiwi#sh ip mroute 233.2.171.1 141.142.64.104
IP Multicast Routing Table
Flags: <cut>
Outgoing interface flags: H - Hardware switched
Timers: Uptime/Expires
Interface state: Interface, Next-Hop or VCD, State/Mode

(141.142.64.104, 233.2.171.1), 6w6d/00:03:26, flags: TA
Incoming interface: Vlan109, RPF nbr 192.5.170.214, Mbgp, RPF-MFD
Outgoing interface list:
  Vlan669, Forward/Sparse, 5d18h/00:02:37, H
```

```
Kiwi#sh ip mroute 233.2.171.1 141.142.64.104 count
IP Multicast Statistics
493 routes using 224398 bytes of memory
71 groups, 5.94 average sources per group
Forwarding Counts: Pkt Count/Pkts per second/Avg Pkt Size/Kbits per sec
Other counts: Total/RPF failed/Other drops(OIF-null, rate-limit etc)

Group: 233.2.171.1, Source count: 123, Group pkt count: 82381322
Source: 141.142.64.104/32, Forwarding: 37847545/9/89/6,Other:33/0/0
```



Engineering Workshops

Trace forwarding state back

1

Juniper:

```
litvanyi@starlight-m10> show multicast route group 233.2.171.1
                           source-prefix 141.142.64.104 extensive

Family: INET
Group          Source prefix    Act Pru NHid  Packets  IfMismtch  Timeout
233.2.171.1    128.55.247.10 /32 A  F  426    5251621    0          360
  Upstream interface: ge-0/0/0.293
  Session name: Static Allocations
  Forwarding rate: 1 kBps (9 pps)
```



Engineering Workshops

Trace forwarding state back

- If you get to a point where the upstream router IS showing it is receiving the packets, but your downstream is not, you need to figure out why those packets are getting lost.
 - ACLs?
 - Broken IGMP snooping switch in the middle?
 - PIM problem?



Trace forwarding state back

- You may work this back to the edge of your area of responsibility, and may have to open a ticket with your upstream to continue the process towards the source. Give them:
 - The active source IP address
 - The group address
 - The circuit / link towards which your router has sent the (S,G) join
 - The fact that you are not receiving packets for that (S,G) on that shared link.



Summary

Gather information

**Verify receiver
interest**

**Verify knowledge of
active source**

**Trace forwarding
state back**



Engineering Workshops

Summary

1

Gather information

- Pick a direction
- Active source and receiver IP addresses
- Group address



Engineering Workshops

Summary

1

**Verify receiver
interest**

- Identify the DR for the receiver.
- Verify the DR knows of interest in that group.
- Check that the DR is not receiving traffic.



Engineering Workshops

Summary

1

Verify knowledge of active source

- Might mean fixing multicast reachability topology or PIM state.
- Probably will involve MSDP SA debugging.



Engineering Workshops

Summary

Trace forwarding state back

- Trace forwarding state from receiver's DR.
- Work towards the actual source.
- Verify reachability, PIM state, and whether traffic is flowing at each step.



A word on troubleshooting performance problems...

1

- Performance problems in multicast inherit virtually all the problems associated with unicast performance issues, which you know how to troubleshoot:
 - packet loss due to congestion.
 - latency/jitter due to queueing, traffic shaping devices, interleaving, etc.
 - duplex problems, cable issues, etc.
- Users often neglect to look at their host performance. Video apps can drive the CPU to where it cannot handle the load.
- It is usually more fruitful to look to the above issues before spending a lot of time looking at timers and such in multicast protocols.



Engineering Workshops

Tools

- **Beacon** <http://dast.nlanr.net/Projects/Beacon/>
 - The beacon is an application to monitor multicast reachability and performance among beacon-group participants. Participants both send and receive on a known group.
 - The results are displayed with receivers on the hosts as receivers on the vertical axis and sources on the horizontal axis.
 - There is a new version out that is RTP-based. By default it runs on a different group than old beacon, and runs at a much slower rate (~1pps vs. ~9pps for old beacon):
<http://dast.nlanr.net/Projects/Beacon/newbeacon/>



Tools

<http://dast.nlanr.net/Projects/Beacon/>

Packet Loss (%)			S0	S1	S2	S3	S4	S5	S6	S7	S8	S9	
R80	[Click for FAQ(2)] agdisplay.chpc.utah.edu	155.101.28.13	NA	0	0	2	0	0	NA	NA	0	0	R80
R81	mcast1.gw.utexas.edu	128.83.6.240	0	0	0	0	0	0	NA	NA	0	0	R81
R82	tulip.as.utk.edu	160.36.8.67	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	R82
R83	[Click for FAQ(2)] ag02.cs.utk.edu	160.36.59.104	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	R83
R84	d-128-208-20-224.dhcp4.washington.edu	128.208.20.224	0	0	0	0	0	0	NA	NA	0	0	R84
R85	mbone-test.cs.wisc.edu	128.105.1.86	0	0	0	0	0	0	NA	NA	0	0	R85
R86	grid-op.trace.wisc.edu	128.104.192.212	0	0	0	2	0	0	NA	NA	0	0	R86
R87	[Click for FAQ(2)] ag-enc.wpi.edu	130.215.128.21	0	0	0	0	0	0	NA	NA	0	0	R87
R88	noc1.wpi.edu	130.215.201.81	7	0	0	10	10	0	NA	NA	0	7	R88
R89	[Click for FAQ(2)] ip-62-54.telcom.wvu.edu	157.182.62.54	NA	0	0	0	0	2	NA	NA	0	NA	R89
Packet Loss (%)			S0	S1	S2	S3	S4	S5	S6	S7	S8	S9	
R90	[Click for FAQ(2)] dsl-agvideo.mcs.anl.gov	140.221.8.157	0	0	0	0	0	0	NA	NA	0	0	R90
R91	[Click for FAQ(2)] lib-video.mcs.anl.gov	140.221.8.53	0	0	0	0	0	0	NA	NA	0	0	R91
R92	ws-video.mcs.anl.gov	140.221.34.1	0	0	0	0	0	0	NA	NA	0	0	R92
R93	[Click for FAQ(2)] micsaudio.er.doe.gov	192.73.213.181	NA	NA	NA	NA	NA	NA	NA	NA	NA	NA	R93
R94	[Click for FAQ(2)] agaudio2.acl.lanl.gov	204.121.50.22	62	70	69	67	65	70	NA	NA	67	71	R94



Engineering Workshops

Tools

- If the beacon is broken, that gives you higher confidence the problem is not just user error or host issues.
- It is sometimes possible to use the beacon as the constantly active source and receiver for debugging.
- However, many times the beacon can be fine yet multicast is broken for a different group.
- It will not catch new/transient problems with source knowledge or state creation (the tree has been built).
- Encourage sites you collaborate with to participate in a beacon group!



Tools

- **Example: GEANT** <http://beaconserver.geant.net:9999>

Time: Sat Feb 08 23:24:51 GMT 2003

Target: 233.81.229.1:56464

Beacons: 12 [details](#)

Page: refresh in 60 seconds

Loss (%)	S0	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11
R0 beacon@62.40.99.107@ws2.lu	0	0	0	0	0	0	0	0	0	0	0	0
R1 beacon@62.40.100.11@ws2.si	0	0	0	0	0	0	0	0	0	0	0	0
R2 beacon@62.40.98.151@ws1.de	0	0	0	0	0	0	0	0	0	0	0	0
R3 beacon@62.40.98.180@ws1.es	0	0	0	0	0	0	0	0	0	0	0	0
R4 beacon@62.40.98.212@ws1.fr	0	0	0	0	0	2	0	0	0	0	0	0
R5 beacon@62.40.98.21@ws2.at	0	0	0	0	0	0	0	0	0	0	0	0
R6 beacon@62.40.99.245@ws2.se	0	0	0	0	0	0	0	0	0	0	0	0
R7 beacon@62.40.98.52@ws1.be	0	0	0	0	0	0	0	0	0	0	0	0
R8 beacon@62.40.100.52@ws1.sk	0	0	0	0	0	0	0	0	0	0	0	0
R9 beacon@62.40.98.85@ws2.ch	0	0	0	0	0	0	0	0	0	0	0	0
R10 beacon@62.40.99.85@ws2.it	0	0	0	0	0	0	0	0	0	0	0	0
R11 beacon@62.40.100.85@ws2.uk	0	0	0	0	0	0	0	0	0	0	0	0



Engineering Workshops

Tools

1

- Some web tools exist to look at peer's routers.
- Again, the Abilene router proxy:
<http://loadrunner.uits.iu.edu/%7Erouterproxy/abilene/>
- Also, some looking-glass pages include multicast information as queries you can run:
<http://www.nordu.net/connectivity/looking-glass/lg.cgi>
- You can get the proxy code free from IU after signing a license agreement. You can freely download the looking glass code and modify it yourself if you would like to make your network visible to others.



Engineering Workshops

Tools

1

Abilene Core Node Router Proxy

A service of the [Abilene NOC](#)

This tool allows you to submit show commands to an Abilene core node router. Select a core node, select and complete the command of your choice, and submit the form; the output of the command will be returned in the lower frame.

Router:

Command:

Response from Router:

Group address	Source address	Peer address	Originator	Flags
233.2.171.1	141.142.64.104	141.142.12.1	141.142.12.1	Accept
		206.220.240.220	141.142.12.1	Reject



Engineering Workshops

Tools

1

The screenshot shows a web browser window titled "NORDUnet Looking Glass - Microsoft Internet Explorer". The address bar shows the URL "http://www.nordu.net/connectivity/looking-glass/lg.cgi". The page content includes the NORDUnet logo, the title "NORDUnet Looking Glass", and a link to "Other looking-glasses". It also provides information about the scripts and a query section.

Query:

☐ bgp ☐ mbgp ☐ ping ☐ Argument:

☐ bgp dampened-paths ☐ mroute ☐ sdr

☐ bgp flap-statistics ☐ mroute summary ☐ trace

☐ bgp summary ☒ msdp

☐ environmental ☐ mtrace

Results of query:

Router: sw-gw.nordu.net
Command: show ip msdp sum

MSDP Peer Address	Status	Summary	Uptime/Down	Reset Count	SA Count	Peer Name
130.225.245.71	Up	1835	7w0d	17	7	lyngby-2.fsknet.lyngby.forskningsnettet.dk
195.54.127.30	Connect	8642	00:00:48	0	0	lo0.cr2.sto1.se.bredband.com
130.242.80.31	Up	1653	4d14h	11	227	stockholm1.sunet.se
62.40.102.33	Up	20965	6w5d	10	4120	lo0.se1.se.geant.net
130.208.17.254	Up	15474	2w5d	11	2	rix-gw.rhnet.is
193.166.255.241	Up	1741	6d04h	20	6	csc0-rtr.funet.fi
80.77.96.5	Up	1239	7w0d	5	180	sl-rp-sto.sprintlink.net
192.108.195.94	Up	1880	2d20h	69	3	dkv37c-GW.Stupi.NET



Engineering Workshops

Tools

- **rtpqual** <ftp://ftp.ee.lbl.gov/rtpqual.c>
 - Simple Multiprotocol Multicast Signal Quality Meter
 - very useful for establishing a receiver (even if the multicast is not using RTP)
 - also useful for finding packet loss problems and whether they are periodic or not



Engineering Workshops

Information Online

- tutorial-style paper at:
<http://multicast.internet2.edu/almeroth.pdf>
- http://www.ncne.nlanr.net/documentation/faq/mcast_eng_faq.html
- <http://dast.nlanr.net/Projects/Beacon/>
- GEANT: <http://www.dante.net/nep/GEANT-MULTICAST/>
links to some troubleshooting docs and monitoring tools
- <ftp://ftpeng.cisco.com/ipmulticast.html>
- <http://www.sprint.net/multicast/faq.html>
- Abilene router proxy:
<http://loadrunner.uits.iu.edu/%7Erouterproxy/abilene/>



Engineering Workshops

Lab 5: MSDP & Inter-domain ASM



Engineering Workshops